## 15.2 StEPS ROSTER PROCESSING

### 15.2.1 Receipt and Initial Validation of Roster Data

Special processing for roster items begins when survey data is received, either electronically or through a form that is keyed. Both processes create the standard data output (SO) file that is used to extract data items flagged as roster items.

- The initial batch update program checks all data items that have the roster flag set to "R" to ensure that a value is present for rtype and rkey1. If either of these required values is missing, an error code will be displayed in the processing log.

- The initial batch update program also checks each ID/ITEM/RTYPE/RKEY1/RKEY2 combination to ensure that it is unique. If it finds duplicate values, it inserts special pseudo codes (i.e., 0999) in the RKEY1 AND RKEY2 variable fields. These display on the Review and Correction screens to alert analysts to the fact that duplicates have been found and must be resolved. (The pseudo codes used are determined by the survey managers and programmers, working with the StEPS programming team.)

- The system then creates a roster FAT file that writes only those ID/ ITEM/ RTYPE/ RKEY1/ RKEY2 combinations with data to the FAT file record. Items without data are eliminated. Roster file records are sorted by ID, RTYPE, REKEY1 and, if specified, RKEY2 for use in editing, computing derived data, imputing data, and displaying data on the Roster Review & Correction screen.

StEPS currently processes roster data at two levels of detail (RKEY1 and RKEY2). Processes to support rosters include review and correction (including an audit trail), batch update, specifying and running edits for roster data, specifying and running simple imputation for roster balance complexes, specifying derived items and estimation for roster values.

### 15.2.2 Review & Correction of Roster Data

Roster data can be accessed from any of the four Review and Correction options:

1. Review and Correction via selection set
2. Review and Correction via data set browse
3. Review and Correction via searches
4. Review and Correction via data analysis

Option 1 enables you to review all the data gathered in a survey period, or to create subsets of this data (for example, all questionnaire IDs from a given state, all IDs for a given SIC). Because this option provides access to the full range of detail review and correction screens, the screen that is displayed when you select Option 1 is called the Review and Correction Main Menu. To

access roster files from this screen, you select the GOTO option on the pmenu, then you select rosters. A submenu under rosters takes you to either the roster data review and correction screens or the roster items audit trail display.

These options are explained in User Manual sections 4.16.

Options 2, 3, and 4 use SAS analysis tools combined with StEPS-developed data search capabilities to examine survey data in detail and correct data in the data set. These options allow you to search for cases meeting certain criteria and view the contents of files. Option 4 uses the SAS INSIGHT program to allow you to produce graphs based on the data analysis.

### To view roster files using Option 2 - Data Set Browse:

Click on the library containing the data files. For the current survey, this is data00. Data01 will hold data from the previous survey period, if it is kept. Phase data from two periods past will be data02. The roster files are:

RRYYYYPP - roster data files for specified stat period
RTYYYYPP - respondent text files associated with roster items for stat period

For details on using this option, see User Manual section 4.1.2.

### To view roster data using Option 3 - Searches:

The following Option 3 preprogrammed searches will display roster items if roster items are specified on the search set-up screen:

Item value/ all cases for an item
Item1 > Item 2
Item1/Item2 > (desired amount)

For details on using this option, see User Manual section 4.1.3.

### To analyze roster data using Option 4 - Data Analysis:

The following selections under Option 4 support analysis of roster data using SAS INSIGHT:

Analyze a selected data set
Select a data set to Analyze
Select 2 data items to analyze
Analyze an Item in current and prior

The data set analysis option is available only if you have previously saved a data set of roster items. For the "comparison of 2 data items" function, the items can be either two roster items or a roster item and a non-roster item. To use the "Item changes by % in current and prior" function, the survey must have data for the item in both the current and prior stat periods.

For details on using these options, see User Manual section 4.1.4.

### 15.2.3 Specifying Edits for Roster Data

Roster edit definitions are only referenced using the Survey rule definition. This is a free-form edit test that validates complex internal relationships within an observation. This test validates that an expression is valid, not that individual data items are valid. Any test can be entered as a Survey Rule edit, as long as it is syntactically correct. Roster edits definitions are stored in the PARMLIB.ROSTED6 file. The data set is named EDIT6 (definitions for Survey Rule tests).

A full discussion of roster edits will be found in User Manual sections 8.4, 8.5, and 8.6.

### 15.2.4 Derived Item Definitions for Roster Items

Specifications for derived items can be created for both parent-level roster items and child-level items, called "second-stage derived items." The reach the screens for these functions, click the DERIVED ITEM DEFINITIONS button on the SURVEY SPECIFICATIONS Menu, then select Option 3 - "Roster Derived Item Definitions."

Detailed instructions for specifying derived definitions for roster items, generating code, and viewing the code are given in User Manual section 7.6.

Derived item definitions are run in batch mode. To specify a batch run for derived roster items, go to the "Run Derived Items" screen. Then select the third option: "f003 – Derive roster items for all IDs." You can access the "Run Derived Items" screen from the RUN PROCESSES button on the StEPS Main Menu or from the GOTO pmenu on the "Derived Item Definitions Menu."

Setting up a batch run for derived items is described in User Manual section 7.3.

### 15.2.5 Roster Item Imputation

If a survey has set up a roster that is a group of related detail items that add to a total, survey processing can use the StEPS roster imputation function. The total may be either an item or a nonnegative integer. The balance complex is said to be out-of-balance when one or more items are missing or when the sum of the detail items does not equal the total. This can happen because one or more items in the complex are missing or because of rounding or keying errors.

Using the roster imputation function, you can specify a series of actions StEPS should take to fill in or adjust data to bring a complex into balance. However, StEPS can provide this functionality only if all values for the balance complex belong to a single case ID (that is, StEPS only imputes roster items for 1-d balance complexes). The system cannot adjust roster items across multiple case IDs. It also cannot adjust detail items when the total is negative.

Roster imputation processing allows you to specify a condition that must exist before imputation takes place. . When a condition has been specified, a balance complex will be processed only when the condition is satisfied.

A detailed discussion of imputation for roster items is given in User Manual section 9.7.